Semi-Implicit and Fully Implicit Shock-Capturing Methods for Hyperbolic Conservation Laws With Stiff Source Terms

H.C. Yee and Judy L. Shinn

December 1986
ERRATA
NASA Technical Memorandum 89415, December 1986

Semi-Implicit and Fully-Implicit Shock-Capturing Methods for Hyperbolic Conservation Laws with Stiff Source Terms
H.C. Yee and Judy L. Shinn

For an updated version of this report, see AIAA-87-1116-CP, Proceedings of the AIAA 8th Computational Fluid Dynamics Conference, June 9-11, 1987, Honolulu, Hawaii.

Pages 6 and 10, equations (2.12c) and (3.1d): The last term should be
\[ \frac{\Delta t}{2} S_{j,k}^{(1)} \]

Page 7: Equation (2.13a) should be
\[ \left( \phi_{j+\frac{1}{2}}^i \right)^S = \frac{1}{2} \left[ \psi(\nu_{j+\frac{1}{2}}^i - (\nu_{j+\frac{1}{2}}^i)^2) \right] (\alpha_{j+\frac{1}{2}}^i - \tilde{Q}_{j+\frac{1}{2}}^i) \]

Page 9: Equation (2.17a), last line should be
\[ H k_2 - \frac{1}{p_e} \sum_{i=1}^{n_s} (a_i^i)^2 \phi^i + v \phi^{n_s+2} + a u k_3 \]

Page 12: Line 6 from top, the word "compressible" should be "compressive".
Semi-Implicit and Fully Implicit Shock-Capturing Methods for Hyperbolic Conservation Laws With Stiff Source Terms

H. C. Yee, Ames Research Center, Moffett Field, California
Judy L. Shinn, Langley Research Center, Hampton, Virginia

December 1986
Semi-Implicit and Fully Implicit Shock-Capturing Methods for Hyperbolic Conservation Laws with Stiff Source Terms

H.C. Yee†
NASA Ames Research Center
Moffett Field, California, 94035
and
Judy L. Shinn‡
NASA Langley Research Center
Hampton, Virginia, 23665-5225

Summary

Some numerical aspects of finite-difference algorithms for nonlinear multidimensional hyperbolic conservation laws with stiff nonhomogenous (source) terms are discussed. If the stiffness is entirely dominated by the source term, a semi-implicit shock-capturing method is proposed provided that the Jacobian of the source terms possesses certain properties. The proposed semi-implicit method can be viewed as a variant of the Bussing and Murman point-implicit scheme with a more appropriate numerical dissipation for the computation of strong shock waves. However, if the stiffness is not solely dominated by the source terms, a fully implicit method would be a better choice. The situation is complicated by problems that are higher than one dimension, and the presence of stiff source terms further complicates the solution procedures for alternating direction implicit (ADI) methods. Several alternatives are discussed. The primary motivation for constructing these schemes was to address thermally and chemically nonequilibrium flows in the hypersonic regime. Due to the unique structure of the eigenvalues and eigenvectors for fluid flows of this type, the computation can be simplified, thus providing a more efficient solution procedure than one might have anticipated. In a companion paper, extension and application of the semi-implicit method to three-dimensional chemically nonequilibrium flows in generalized coordinates will be presented. A preliminary numerical computation for a five-species reacting flow is shown to be oscillation-free around the shock.

I. Motivation and Objectives

The second-order total variation diminishing (TVD) type of schemes [1-5] originally designed for nonlinear scalar hyperbolic conservation laws or constant coefficient hyperbolic systems has been proven to be applicable to many multidimensional fluid dynamics problems (especially for a perfect gas). See for example, references [6-11]. In particular, the explicit and implicit

†Research Scientist, Computational Fluid Dynamics Branch
‡Aero-Space Technologist, Aerothermodynamics Branch, Senior Member AIAA
TVD schemes originally developed by Harten [1], Roe [12] and Davis [13], and modified and generalized by Yee [2] are rather simple in structure and require the least computational effort among the class of TVD methods. For the two-dimensional, compressible, Euler equations for a perfect gas, these explicit TVD methods require approximately twice the CPU time as the classical Lax-Wendroff and MacCormack schemes. The gain in robustness and elimination of spurious oscillation by TVD schemes often justifies the extra computation. On the other hand, straightforward extension of these methods to include stiff source terms for large systems is prohibitively expensive in terms of efficiency. There has been little work in efficiently generalizing TVD methods to large systems with stiff source terms. Here some numerical aspects and new procedures based on existing TVD properties, semi-implicit methods, relaxation methods, and the choice of the physical equation set in treating nonequilibrium flows will be discussed.

**Stiffness of Source Terms:** Treating stiff problems implicitly is a common practice for the integration of stiff ordinary differential equations [14]. To further speed up the process, Kreiss [15] proposed splitting the equation into stiff and non-stiff terms and applying a special initial condition to damp out the fast wave initially. Guerra and Gustafsson [16] extended these ideas by applying a semi-implicit method to time-dependent hyperbolic problems with varying time scales. Their technique is to treat the stiff part implicitly and at the same time damp out the fast wave, leaving a more manageable system. Their methodology works well for nearly incompressible flows or transonic flows. For general hypersonic applications, the situation is more complicated. However, if the stiffness is dominated by the source terms and steady-state solutions are especially desired, one can still use this idea for hypersonic study by treating the source term implicitly and the rest explicitly.

In the area of both thermally and chemically nonequilibrium flows, the terms responsible for the stiffness are the source terms. Kee and Miller [17] used a first-order spatial differencing together with a fully implicit time-marching scheme in handling combustion-related problems. In order to accurately capture complex flow fields, higher-order spatial differencing is more desirable. The most common types of higher-order spatial differencing are the second- and fourth-order central differences. Bussing and Murman [18] have illustrated that taking second-order central difference in space and treating the source term implicitly will essentially rescale the equation in time so that all events occur on a similar pseudo-time scale for steady-state chemical reaction calculations. Similarly, other researchers such as [11,19-23] have utilized some kind of implicit treatment for stiff problems.

**Classical vs Modern Shock-Capturing Schemes:** Most of the methods discussed above used classical shock-capturing methods. They only exhibit accurate results for smooth or weak shock solutions and are not robust enough for hypersonic computations. This is typical of classical shock-capturing techniques that result in oscillatory solutions across the discontinuities and/or nonlinear instability, while the modern shock-capturing techniques such as the upwind and symmetric TVD schemes of Harten [1], Yee [2], Roe [3,12], Osher and Chakravarthy [4], and Jameson and Lax [5] do not possess these deficiencies. In particular, the symmetric TVD schemes appears to be more attractive than the upwind TVD schemes since the symmetric schemes have a smaller operation count and are less sensitive to numerical boundary-condition treatments [7]. An example is the current enhancement work to program LAURA by Gnoffo et al. [11] which has resulted in a large savings in TVD operation steps.
Coupling or Uncoupling the Species Equation from the Fluid Equations: In the usage of modern shock-capturing methods like the TVD type of schemes for the chemically reacting flows, Carofano [10] was the first to introduce the formulism that enabled full coupling in Harten's explicit TVD scheme for a two-species, two-dimensional unsteady flow in Cartesian coordinates. To avoid solving a large system, Gnoffo and McCandless [19], uncoupled the species equations from the fluid dynamics equations and solved these two sets of systems of nonlinear partial differential equations in a time-lag fashion (loosely coupled method) by using a point-relaxation technique with a second-order symmetric TVD scheme of Yee [2] and an upwind TVD scheme of Osher and Chakravarthy [4]. Eberhardt and Brown [24] attempted to use the eigenvalues and eigenvectors of the fluid dynamics equations alone to obtain a “fully coupled” first-order explicit TVD scheme for a one-dimensional flow. The result of Eberhardt and Brown showed excessive smearing at the shock when compared to the true fully coupled explicit TVD result. Their motivation for designing such a coupling procedure was to optimize the operation count by avoiding multiplication of large matrices. However, as will be demonstrated in the present work, if one makes use of the unique structure of eigenvectors and eigenvalues for fluid flow of this type, the fully coupled formulation can be simplified even for a large number of species, thus providing a more efficient solution procedure than one might have anticipated. Moreover, using the eigenvalues and eigenvectors for the fully coupled equation set allows one to have the freedom of controlling the proper amount of numerical dissipation for the individual waves [7]. In particular, for the two-dimensional chemically reacting flows, the number of linear waves is \( n_s + 1 \) where \( n_s \) is the number of species. Note that in order to capture contact discontinuities accurately, it is very important to apply the proper amount of numerical dissipation to the linear waves.

Choosing Equation Set: In the present formulation for the chemically reacting flows, the global continuity equation is replaced by all the species continuity equations, instead of keeping the global continuity equation minus one species equation. Here the redundancy in species equations due to elemental conservation is assumed to be eliminated in advance. Our study shows that the latter approach has two deficiencies, namely: (a) the solution of the absent species will suffer a loss of numerical accuracy due to arithmetic subtraction since it is equal to the total density minus the rest of the species density, and (b) the eigenvectors and its inverses, which are required for the proposed schemes, are slightly more complicated than in the former formulation.

Proposing New Schemes: The goal of the present work is to construct second-order high-resolution schemes such that they are appropriate for thermally and chemically nonequilibrium flows in the hypersonic regime, and, at the same time, they are efficient and simple to program and easy to implement into existing computer codes. The approach here is to improve the existing technology by combining the desirable features from a few existing numerical methods into a single algorithm and, most importantly, take into account the specific characteristics of the particular physical problem.

Two types of schemes are proposed. If the stiffness is entirely dominated by the source term, a semi-implicit TVD type of shock-capturing method is proposed for both time-accurate and steady-state calculations provided that the Jacobian of the source terms possesses certain properties. The proposed semi-implicit scheme can be viewed as a variant of the Bussing and Murman point-implicit predictor-corrector scheme [18] with a more appropriate numerical
dissipation in the computation of strong shock waves in the hypersonic regime and a speed up in convergence rate for steady-state applications. The predictor-corrector scheme of Bussing and Murman in turn is the explicit MacCormack scheme with the source term treated implicitly. The main advantage of the predictor-corrector formulation over most other one-step second-order TVD schemes is that, for multidimensional problems, one can incorporate the source terms more efficiently than in other formulisms.

However, if the stiffness is not solely dominated by the source terms (e.g., stiffness due to highly irregular grid and/or viscous flows), a fully implicit method would be a better choice. The situation is complicated by problems that are higher than one dimension, and the presence of stiff source terms further complicates the solution procedures for alternating direction implicit (ADI) methods. In fact, there seems to be no straightforward way of treating general stiff source terms implicitly with ADI procedures. Several alternatives will be discussed. The proposed fully implicit relaxation algorithm can be viewed as a variant of a fully coupled form of the algorithm proposed by Gnooffo and McCandless [19].

Due to space limitations, the material is divided into two separate papers. The mathematical aspects and the derivation of the numerical methods for solving the two-dimensional hypersonic nonequilibrium flow in cartesian coordinates are given in this paper, while the extension of the semi-implicit method to the three-dimensional, fully coupled chemically reacting viscous flow in generalized coordinates will be presented in a companion paper [25]. In the following sections, the proposed numerical methods for both the semi-implicit and fully implicit shock-capturing schemes will be briefly described. A few test examples will also be included to support the performance of the proposed schemes.

II. The Governing Equations and the Basic Schemes

In this section, the predictor-corrector form of the basic explicit TVD scheme and an example to illustrate the performance of the scheme are given. Also, all the necessary terms which are required for the basic TVD scheme for the compressible inviscid chemically reacting flow equations are derived.

2.1 The Governing Equations

Consider a two-dimensional system of nonhomogenous hyperbolic conservation laws

$$\frac{\partial U}{\partial t} + \frac{\partial F(U)}{\partial x} + \frac{\partial G(U)}{\partial y} = S(U). \quad (2.1)$$

Here $U$, $F(U)$, $G(U)$, and $S(U)$ are column vectors of $k$ components. Let $A = \frac{\partial F}{\partial U}$ and $B = \frac{\partial G}{\partial U}$, with $(a_1^x, a_2^x, \ldots, a_k^x)$ and $(a_1^y, a_2^y, \ldots, a_k^y)$ being the eigenvalues of $A$ and $B$. Denote $R_x$ and $R_y$ as the matrices whose columns are eigenvectors of $A$ and $B$, and denote $R_x^{-1}$ and $R_y^{-1}$ as the inverses of $R_x$ and $R_y$. In the case of the compressible inviscid flow equations with chemical reactions where the global continuity equation is replaced by all the species continuity equations,
\[ U = \begin{bmatrix} \rho^1 \\ \rho^2 \\ \vdots \\ \rho^{n_s} \\ m \\ n \\ e \end{bmatrix}; \quad F(U) = \begin{bmatrix} c^1 m \\ c^2 m \\ \vdots \\ c^{n_s} m \\ \frac{m^2}{\rho} + p \\ \frac{m n}{\rho} \\ (e + p) \frac{m}{\rho} \end{bmatrix}; \quad G(U) = \begin{bmatrix} c^1 n \\ c^2 n \\ \vdots \\ c^{n_s} n \\ \frac{m n}{\rho} \\ \frac{n^2}{\rho} + p \\ (e + p) \frac{n}{\rho} \end{bmatrix}; \quad S(U) = \begin{bmatrix} s^1 \\ s^2 \\ \vdots \\ s^{n_s} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad (2.2) \]

Here \( m = \rho u, n = \rho v, \) and \( s^i \) represents the production of species from chemical reactions. The variables are the velocity components \( u \) and \( v, \) the pressure \( p, \) the energy per unit volume \( e, \) and the density of the \( i \)th species \( \rho^i \) with \( \rho = \sum_{i=1}^{n_s} \rho^i, \) \( c^i \rho^i = \rho^i, \) where \( n_s \) is the number of species in the model and \( c^i \) is the species mass fraction. Equation (2.2) assumes all of \( \rho^i \) are linearly independent. If due to elemental conservation, \( n_s \) should be the number of species after the redundancy in the linearly dependent species is eliminated.

The eigenvalues of \( A \) and \( B \) are

\[ (a_x^1, \ldots, a_x^{n_s+3}) = (u, \ldots, u, u + a, u, u - a) \quad (2.3a) \]

\[ (a_y^1, \ldots, a_y^{n_s+3}) = (v, \ldots, v, v + a, v, v - a). \quad (2.3b) \]

Here the frozen sound speed \( a \) is

\[ a^2 = p^i + p_e(H - u^2 - v^2) \quad (2.4) \]

with

\[ p^i = \sum_{i=1}^{n_s} c^i p^{i}, \quad (2.5) \]

\[ \sum_{i=1}^{n_s} c^i = 1 \quad (2.6) \]

\[ \frac{\partial p^i}{\partial \rho^i} \bigg|_{m,n,e,\ldots} \quad (2.7) \]

\[ \frac{\partial p}{\partial e} \bigg|_{m,n,\rho^i,\ldots} \quad (2.8) \]

\[ H = \frac{e + p}{\rho} \quad (2.9) \]

Let the grid spacing be denoted by \( \Delta x \) and \( \Delta y \) such that \( x = j \Delta x \) and \( y = k \Delta y, \) and let the time step be denoted by \( \Delta t \) such that \( t = n \Delta t \) where the superscript "n" is used for the
time index and should not be confused with the \( n = \rho u \) in equation (2.2). Let \( a_l^{j+\frac{1}{2}}, R_x^{j+\frac{1}{2}}, R_x^{-1} \) denote the quantities \( a_l^i, R_x, R_x^{-1} \) evaluated at some symmetric average of \( U_{j,k} \) and \( U_{j+1,k} \). Similarly, let \( a_l^{k+\frac{1}{2}}, R_y^{k+\frac{1}{2}}, R_y^{-1} \) denote the quantities \( a_l^i, R_y, R_y^{-1} \) evaluated at some symmetric average of \( U_{j,k} \) and \( U_{j,k+1} \). In the case of the chemically reacting flows, \( a_l^{j+\frac{1}{2}}, R_x^{j+\frac{1}{2}}, R_x^{-1}, a_l^{k+\frac{1}{2}}, R_y^{k+\frac{1}{2}}, R_y^{-1} \) are defined similar to the ones used by Huang [26] and Carofano [10]. Also, the commonly used Roe’s average is no longer valid for a nonperfect gas [27].

Define

\[
\alpha_{j+\frac{1}{2}} = R_x^{-1} (U_{j+1,k} - U_{j,k})
\]

(2.10)

as the difference (or the jump) of the characteristic variables in the locally \( x \)-direction, and define

\[
\alpha_{k+\frac{1}{2}} = R_y^{-1} (U_{j,k+1} - U_{j,k})
\]

(2.11)

as the difference of the characteristic variables in the locally \( y \)-direction. For thermally and chemically nonequilibrium flows, the eigenvalues and eigenvectors have a similar structure. For the two-dimensional system (2.1), if \( nt \) is the number of thermal energy variables, then the eigenvalues in the \( x \)-direction will have \((ns + nt + 1)\) \( \alpha \) characteristics plus \( u + a \) and \( u - a \) characteristics. Here the values \( \alpha \) will reflect the added thermal energy variables.

2.2 Explicit Preditor-Corrector TVD Scheme

With the above notation, a formal extension of the scalar explicit second-order TVD method [2] in predictor-corrector form via the local characteristic approach for the nonlinear hyperbolic system (2.1) can be written as

\[
\Delta U_{j,k}^{(1)} = -\frac{\Delta t}{\Delta x} \left( F^n_{j,k} - F^n_{j-1,k} \right) - \frac{\Delta t}{\Delta y} \left( G^n_{j,k+1} - G^n_{j,k} \right) + \Delta t S^n_{j,k}
\]

(2.12a)

\[
U_{j,k}^{(1)} = U_{j,k}^{(1)} + U_{j,k}^n
\]

(2.12b)

\[
\Delta U_{j,k}^{(2)} = \frac{1}{2} \left\{ -\Delta U_{j,k}^{(1)} - \frac{\Delta t}{\Delta x} \left[ F^{(1)}_{j+1,k} - F^{(1)}_{j,k} \right] - \frac{\Delta t}{\Delta y} \left[ G^{(1)}_{j,k+1} - G^{(1)}_{j,k} \right] \right\} + \Delta t S_{j,k}^{(1)}
\]

(2.12c)

\[
U_{j,k}^{(2)} = U_{j,k}^{(2)} + U_{j,k}^{(1)}
\]

(2.12d)

\[
U_{j,k}^{n+1} = U_{j,k}^{(2)} + \left[ R_x^{(2)} \Phi_{j+\frac{1}{2}}^{(2)} - R_x^{(2)} \Phi_{j-\frac{1}{2}}^{(2)} \right] + \left[ R_y^{(2)} \Phi_{j+\frac{1}{2}}^{(2)} - R_y^{(2)} \Phi_{j-\frac{1}{2}}^{(2)} \right]
\]

(2.12e)

Here the superscripts “(1) and (2)” designate the values of the function evaluated at the intermediate solutions \( U^{(1)} \) and \( U^{(2)} \). The elements of the vector \( \Phi_{j+\frac{1}{2}} \), denoted by \( (\phi_{j+\frac{1}{2}}) \) with \( l = 1, ..., k \) for a symmetric TVD scheme, are
where $\alpha^t_{j+\frac{1}{2}}$ are elements of (2.10) and $\alpha^t_{j+\frac{1}{2}}$ are eigenvalues in the $x$-direction. The function $\psi$ is

$$\psi(z) = \begin{cases} |z| & |z| \geq \epsilon \\ (z^2 + \epsilon^2)/2\epsilon & |z| < \epsilon \end{cases}$$

(2.13c)

Here $\psi(z)$ in (2.13c) is an entropy correction to $|z|$, where $\epsilon$ is a small positive parameter (see reference 1281 for a formula for $\epsilon$). In practical calculations, $0.05 \leq \epsilon \leq 0.2$ is a commonly used range. Examples of the 'limiter' function $\tilde{Q}_j^{t+\frac{1}{2}}$ can be expressed as

$$\tilde{Q}_j^{t+\frac{1}{2}} = \min\text{mod}(\alpha^t_{j-\frac{1}{2}}, \alpha^t_{j+\frac{1}{2}}) + \min\text{mod}(\alpha^t_{j+\frac{1}{2}}, \alpha^t_{j+\frac{3}{2}}) - \alpha^t_{j+\frac{1}{2}}$$

(2.13d)

$$\tilde{Q}_j^{t+\frac{1}{2}} = \min\text{mod}(\alpha^t_{j-\frac{1}{2}}, \alpha^t_{j+\frac{1}{2}}, \alpha^t_{j+\frac{3}{2}})$$

(2.13e)

$$\tilde{Q}_j^{t+\frac{1}{2}} = \min\text{mod}[2\alpha^t_{j-\frac{1}{2}}, 2\alpha^t_{j+\frac{1}{2}}, 2\alpha^t_{j+\frac{3}{2}}, \frac{1}{2}(\alpha^t_{j-\frac{1}{2}} + \alpha^t_{j+\frac{3}{2}})]$$

(2.13f)

The minmod function of a list of arguments is equal to the smallest number in absolute value if the list of arguments is of the same sign, or is equal to zero if any arguments are of opposite sign.

Here, the use of the nomenclature “TVD scheme” pertains to schemes satisfying the TVD sufficient conditions for one-dimensional nonlinear scalar hyperbolic conservation laws and/or constant coefficient hyperbolic systems. The second-order scheme (2.12) belongs to the class of symmetric TVD methods as discussed in reference [121]. By defining a more complex $\phi^t_{j+\frac{1}{2}}$, (2.12) can be made upwind weighted and would belong to the class of upwind TVD schemes. The derivation is straightforward and will not be given here.

The formulation of this scheme is purposely broken into two parts, namely, the predictor-corrector step of the MacCormack explicit scheme, and an appropriate conservative dissipation term designed in such a way that the final scheme is TVD for the one-dimensional constant coefficient case. This predictor-corrector TVD method is sometimes referred as the “TVD MacCormack” scheme. It is a slight modification of Roe’s one-step TVD Lax-Wendroff scheme. If one sets $\tilde{Q}$ to be equation (2.13d) and $\psi(z) = |z|$, the scalar scheme is the same as described in Davis [13] and Kwong [29]. The reason for choosing the predictor-corrector step instead of the one-step Lax-Wendroff formulation is that the predictor-corrector form provides a natural and efficient inclusion of the source terms for multidimensional problems.

Another numerical aspect is that the method of extending scalar TVD schemes to nonlinear systems of hyperbolic conservation laws (2.1) in equation (2.12) by using the eigenvalues and eigenvectors of $\frac{\partial F}{\partial U}$ and $\frac{\partial G}{\partial U}$, is sometimes referred to as the local characteristic approach and is a variant of Roe’s linear Riemann solver [30]. The advantages of this approach as opposed to Davis’s simplified approach [13] to systems are that (a) the current approach in
effect uses scalar schemes on each characteristic field; (b) the limiter used need not be the same for each field; e.g., one can use a more compressive limiter for the linear fields and use a less compressive limiter for the nonlinear fields; see references [2,7] for some numerical examples; (c) one can even use different schemes for different fields; (d) it is more efficient than the exact or approximate Riemann solvers; and (e) it provides a natural way to linearize the implicit TVD schemes [8,31]. For the hyperbolic system (2.1)-(2.9), the characteristic fields consist of two nonlinear fields “\(u \pm a\)” and \((ns + 1)\) linear fields of the same value “\(u\)” The contact discontinuities are associated with the linear fields. It has been shown [2,7,9] that the two different fields required different amounts of numerical dissipation (i.e., different limiters). Often the limiters that are designed for the linear field might give spurious oscillation or nonphysical solutions for the nonlinear field.

To illustrate the performance of this scheme, figures 1 and 2 show a computation of a moving planar shock at an incident Mach number \(M_S = 2\) interacting with an obstacle using the scheme (2.12) together with the limiter (2.13f). The calculation is for a perfect gas where \(S = 0\). The grid-generation method used for this study is a generalized Schwartz-Christoffel transformation. Figure 2 shows eight sequential stages of the diffraction process. The results indicate that the explicit scheme (2.12) is capable of capturing even the subtle flow discontinuities such as the slipstreams. A detailed description of the computation is reported by Young and Yee [32]. The results of references [2,7,9,32] give a strong indication that algorithm (2.12) should be a good candidate for solving the hypersonic nonequilibrium flows, and further inspired the authors to improve scheme (2.12) for more efficient solution procedures.

2.3 More Efficient Solution Procedures

The extra computation in (2.12) compared with the classical central difference shock-capturing scheme such as the Lax-Wendroff method is due to the vectors \((R\Phi)_{j\pm \frac{1}{2}}\). At first glance, the vectors \(\alpha_{j\pm \frac{1}{2}}\) and \((R\Phi)_{j\pm \frac{1}{2}}\) involve matrix and vector multiplication of dimension \(ns + 3\) for every grid point, and thus tend to discourage their adaption to problems other than ideal gas flows. Researchers such as Gnoflo and McCandless [19], and Eberhardt and Brown [24] were motivated to pursue other avenues to solve the complicated chemically reacting flow problems. However, as will be demonstrated in this subsection, if one makes use of the unique structure of the eigenvectors and eigenvalues for fluid flow of this type, the fully coupled formulation can be simplified even for a large number of species, and thus be a viable approach.

With straightforward manipulations, the operation for scheme (2.12) can be simplified tremendously. The vector \(\alpha\) in equation (2.10), for example, can be expressed as
\[
\alpha = \begin{bmatrix}
\Delta \rho^1 - c^1 aa \\
\Delta \rho^2 - c^2 aa \\
\vdots \\
\Delta \rho^n - c^n aa \\
\frac{1}{2} (aa - u^2 bb + \Delta m) \\
- vbb + \Delta n \\
\frac{1}{2} (aa + u^2 bb - \Delta m)
\end{bmatrix}
\] (2.14)

with

\[
aa = \frac{1}{a^2} \left[ \sum_{i=1}^{n_s} p_{ii} \Delta \rho^i - p_e (u\Delta m + v\Delta n - \Delta e) \right]
\] (2.15)

\[
bb = \sum_{i=1}^{n_s} \Delta \rho^i ,
\] (2.16)

here, for example, \((\Delta \rho^j)_{j+\frac{1}{2}} = \rho^j_{j+1} - \rho^j_j\) and it is understood that the spatial indices in (2.14)-(2.16) are at \((j + \frac{1}{2}, k)\). Similarly, \(R\Phi\) also has a very simple form. For instance, the \(R\Phi\) associated with the \(x\)-direction flux can be expressed as

\[
R\Phi = \begin{bmatrix}
\phi^1 + c^1 k_1 \\
\phi^2 + c^2 k_1 \\
\vdots \\
\phi^{n_s} + c^{n_s} k_1 \\
u k_2 + ak_3 \\
v k_2 + \phi^{n_s+2} \\
H k_2 - \frac{1}{p_e} \sum_{i=1}^{n_s} (a^i)^2 \phi^i + v k_3
\end{bmatrix}
\] (2.17a)

with

\[
k_1 = \phi^{n_s+1} + \phi^{n_s+3}
\] (2.17b)

\[
k_2 = \sum_{i=1,i\neq n_s+2}^{n_s+3} \phi^i
\] (2.17c)

\[
k_3 = \phi^{n_s+1} - \phi^{n_s+3}
\] (2.17d)

Here the spatial indices on (2.17) are at \((j + \frac{1}{2}, k)\). As one can see, the terms in equations (2.14) and (2.17) due to the species equations are simple and do not require many operations. Therefore, the increase in the number of species equations is not as “CPU” intensive as one might have anticipated.
III. A Semi-implicit Predictor-Corrector TVD Scheme

The explicit TVD scheme (2.12) can be used for either time-accurate or steady-state calculation. It is second-order-accurate in time and space. However, if the source term is stiff, the restriction in the time step due to stability requirement is prohibitively small and (2.12) is not practical, especially for steady-state applications. In this section, a semi-implicit method is proposed. Another alternative is a fully implicit method. The basic implicit scheme and the related difficulty in extending the implicit method to higher dimensions with stiff source terms will be discussed in section IV.

If one follows the same idea as Bussing and Murman [18] in treating the source term implicitly, a semi-implicit predictor-corrector TVD scheme can easily be obtained. It can be written as a one-parameter family of time-differencing schemes for the source term; i.e., the following formulation includes scheme (2.12). The proposed scheme can be written as

\[
D^n_{j,k} U_{j,k}^{(1)} = -\frac{\Delta t}{\Delta x} \left( P^n_{j,k} - P^n_{j-1,k} \right) - \frac{\Delta t}{\Delta y} \left( G^n_{j,k+1} - G^n_{j,k} \right) + \Delta t S^n_{j,k}
\]  

\[
D = \left( I - \Delta t \theta \frac{\partial S}{\partial U} \right)
\]

\[
U_{j,k}^{(1)} = \Delta U_{j,k}^{(1)} + U_{j,k}^n
\]

\[
D^{(1)}_{j,k} U_{j,k}^{(2)} = \frac{1}{2} \left\{ -\Delta U_{j,k}^{(1)} - \frac{\Delta t}{\Delta x} \left[ F_{j+1,k}^{(1)} - F_{j,k}^{(1)} \right] - \frac{\Delta t}{\Delta y} \left[ G_{j,k}^{(1)} - G_{j,k-1}^{(1)} \right] \right\} + \Delta t S_{j,k}^{(1)}
\]

\[
U_{j,k}^{(2)} = \Delta U_{j,k}^{(2)} + U_{j,k}^{(1)}
\]

\[
U_{j,k}^{n+1} = U_{j,k}^{(2)} + \left[ R_{j+\frac{1}{2}}^{(2)} \Phi_{j+\frac{1}{2}}^{(2)} - R_{j-\frac{1}{2}}^{(2)} \Phi_{j-\frac{1}{2}}^{(2)} \right] + \left[ R_{k+\frac{1}{2}}^{(2)} \Phi_{k+\frac{1}{2}}^{(2)} - R_{k-\frac{1}{2}}^{(2)} \Phi_{k-\frac{1}{2}}^{(2)} \right].
\]

Here, \( D \) is assumed to be invertible; i.e., only the type of source terms with its Jacobian such that \( D \) is invertible at each grid point is permissible. The parameter \( \theta \) is in the range \( 0 \leq \theta \leq 1 \). For \( \theta \neq 0 \), the source terms are treated implicitly. If \( \theta = 1 \), the time-differencing for the source term is first-order and (3.1) is suitable for steady-state calculations. But if \( \theta = 1/2 \), the time-differencing is second-order and (3.1) is suitable for time-accurate calculations.

One can simplify equation (3.1) by partitioning the vectors \( U, F, G, S, D \) in equation (2.2) as follows:

\[
U = \begin{bmatrix} U^I \\ U^{II} \end{bmatrix}; \quad U^I = \begin{bmatrix} u^1 \\ \vdots \\ u^n \end{bmatrix}; \quad U^{II} = \begin{bmatrix} u^{n+1} \\ u^{n+2} \\ u^{n+3} \end{bmatrix}
\]  

(3.2a)
Here $D^{21}$ is a null matrix and $D^{22}$ is an identity matrix. With the above definitions, the scheme can be greatly simplified. The procedures are as follows: taking the predictor step, for example, one first solves for $(\Delta U^{II})^{(1)}_{j,k}$ by

$$
(\Delta U^{II})^{(1)}_{j,k} = - \frac{\Delta t}{\Delta x} \left[ (F^{II})^n_{j,k} - (F^{II})^n_{j-1,k} \right] - \frac{\Delta t}{\Delta y} \left[ (G^{II})^n_{j,k+1} - (G^{II})^n_{j,k} \right]
$$

then for $\Delta U^I$ by

$$
(D^{11})^n_{j,k} (\Delta U^I)_{j,k} = (\text{r.h.s})^I_{j,k} - D^{12}_{j,k} (\Delta U^{II})^{(1)}_{j,k},
$$

where $(\text{r.h.s})^I$ is the right-hand side of $(3.3a)$ with all the indices “II” replaced by “I” and with the term $\Delta t(S^n_{j,k})$ added. In other words, one only has to invert the $D^{11}$ matrix of dimension $(ns, ns)$ instead of $(ns + 3, ns + 3)$. Similarly, one can simplify the corrector step in the same way. Or, to explain it in another way, one solves the fluid equations

$$
\frac{\partial U^{II}}{\partial t} + \frac{\partial F^{II}(U)}{\partial x} + \frac{\partial G^{II}(U)}{\partial y} = 0
$$

explicitly, then uses the result to solve the species equations

$$
\frac{\partial U^I}{\partial t} + \frac{\partial F^I(U)}{\partial x} + \frac{\partial G^I(U)}{\partial y} = S^I
$$

explicitly, with the exception that the chemical reaction terms are treated implicitly in $(3.4b)$. In the case where $S^{II}$ is not a null vector and it is not stiff, equation $(3.3)$ is still applicable, except one has to add the term $\Delta t(S^n_{j,k})$ to the right-hand side of $(3.3a)$. In the steady-state calculations where body fitted coordinates are used, one can further speed up the convergence rate by using a local time-stepping approach [33,34].

Figures 3-5 shows a preliminary test result for a three-dimensional, five-species reacting flow using the semi-implicit TVD scheme $(3.1)$ together with $(2.13d)$ as compared with an existing classical shock-capturing method which supplies numerical dissipation linearly [23]. The numerical result is shown to be oscillation-free around the shock while the time spent per iteration is approximately double when compared with the method used in [23]. The
configuration of the numerical experiment is shown in figure 3. The inflow conditions are \( p = 1 \text{ atm} \), the temperature \( T = 1200 \text{ K} \), and the Mach number \( M = 4 \) for a premixed air and hydrogen fuel. The species considered are \( H_2, O_2, OH, H_2O, \) and \( N_2 \), with \( N_2 \) being inert. The so-called \( \text{"global two-step chemistry model\"} \) for combustion is used.

This result is very preliminary and the shock can be further sharpened with more compressible and/or upwind weighted limiters. The resulting research computer code for the test problem was developed based on an existing code developed by A. Kumar of NASA Langley, and K. Uenishi of Vigyan Research Inc. Their original code uses a semi-implicit classical shock-capturing method. A detailed description of the numerical experiments and the extension of (3.1) to three-dimensional, chemically reacting flows in generalized coordinates is given in the companion paper [25].

IV. A Fully Implicit TVD Method

Another type of shock-capturing scheme that the authors are familiar with is a one-parameter family of explicit and implicit TVD type of schemes [2,31]. For the nonequilibrium equation (2.1), these schemes can be written as

\[
U_{j,k}^{n+1} + \theta \left[ \frac{\Delta t}{\Delta x} \left( \tilde{F}_{j+\frac{1}{2},k}^{n+1} - \tilde{F}_{j-\frac{1}{2},k}^{n+1} \right) + \frac{\Delta t}{\Delta y} \left( \tilde{G}_{j,k+\frac{1}{2}}^{n+1} - \tilde{G}_{j,k-\frac{1}{2}}^{n+1} \right) \right] - \Delta t S_{j,k}^{n+1} = U_{j,k}^n - (1 - \theta) \left[ \frac{\Delta t}{\Delta x} \left( \tilde{F}_{j+\frac{1}{2},k}^{n} - \tilde{F}_{j-\frac{1}{2},k}^{n} \right) + \frac{\Delta t}{\Delta y} \left( \tilde{G}_{j,k+\frac{1}{2}}^{n} - \tilde{G}_{j,k-\frac{1}{2}}^{n} \right) - \Delta t S_{j,k}^{n} \right]. \tag{4.1a}
\]

Here \( \theta \) has the same meaning as before, except the scheme is now fully implicit for \( \theta \neq 0 \).

The numerical flux function \( \tilde{F}_{j+\frac{1}{2},k} \) for both the upwind and symmetric TVD schemes [2,8] can be expressed as

\[
\tilde{F}_{j+\frac{1}{2},k} = \frac{1}{2} \left( F_{j,k} + F_{j+1,k} + R_{j+\frac{1}{2}} \Phi_{j+\frac{1}{2}} \right). \tag{4.1b}
\]

Similarly, one can define the numerical flux \( \tilde{G}_{j,k+\frac{1}{2}} \) in this manner. The elements of the \( \Phi_{j+\frac{1}{2}} \) denoted by \( (\phi_{j+\frac{1}{2}}^{(l)})^S \) for a general second-order symmetric TVD scheme are

\[
(\phi_{j+\frac{1}{2}}^{(l)})^S = -\psi(a_{j+\frac{1}{2}}^{(l)})(a_{j+\frac{1}{2}}^{(l)} - \tilde{Q}_{j+\frac{1}{2}}^{(l)}), \tag{4.2}
\]

where \( a_{j+\frac{1}{2}}^{(l)} \) are elements of (2.10). The function \( \psi \) is (2.13c). The \( \text{"limiter\"} \) function \( \tilde{Q}_{j+\frac{1}{2}}^{(l)} \) can be the same as that in equations (2.13d)-(2.13f).

The elements of the \( \Phi_{j+\frac{1}{2}} \) denoted by \( (\phi_{j+\frac{1}{2}}^{(l)})^U \) for a second-order upwind TVD scheme, originally developed by Harten [1] and later modified and generalized by the first author [7,31], are

\[
(\phi_{j+\frac{1}{2}}^{(l)})^U = \frac{1}{2} \psi(a_{j+\frac{1}{2}}^{(l)})(g_{j+1}^{(l)} + g_{j}^{(l)}) - \psi(a_{j+\frac{1}{2}}^{(l)} + g_{j+\frac{1}{2}}^{(l)})a_{j+\frac{1}{2}}^{(l)}. \tag{4.3a}
\]

and
Examples of the “limiter” function $g_j^l$ can be expressed as

$$g_j^l = \min(\alpha_j^l - \frac{1}{2}, \alpha_j^{l+1})$$

$$g_j^l = \left(\alpha_j^{l+1} \alpha_j^l - \frac{1}{2} + |\alpha_j^{l+1} \alpha_j^l - \frac{1}{2}|\right) / \left(\alpha_j^{l+1} + \alpha_j^l - \frac{1}{2}\right)$$

$$g_j^l = S \cdot \max \left[0, \min(2 |\alpha_j^{l+1} |, S \cdot \alpha_j^l - \frac{1}{2}), \min(2 |\alpha_j^{l+1} |, 2S \cdot \alpha_j^l - \frac{1}{2})\right]; S = \text{sgn}(\alpha_j^{l+1}).$$

### 4.1 A Conservative Linearized Form For Steady-State Applications

To solve for $U^{n+1}$ in (4.1), one normally needs to solve a set of nonlinear algebraic equations iteratively. One way to avoid this is to linearized the implicit operator, and solve the linearized form by other means. Following the same procedure as in Yee [31], a conservative linearized form of (4.1) can be written as

$$\left\{ I + \theta \left[ \frac{\Delta t}{\Delta x} \left( H_{j+\frac{1}{2},k}^x - H_{j-\frac{1}{2},k}^x \right) + \frac{\Delta t}{\Delta y} \left( H_{j,k+\frac{1}{2}}^y - H_{j,k-\frac{1}{2}}^y \right) - \Delta t D_{j,k}^n \right] \right\} E = U^{n+1} - U^n$$

where

$$H_{j+\frac{1}{2},k}^x = \frac{1}{2} \left( A_{j+1,k} - \Omega_{j+\frac{1}{2},k}^x \right)^n$$

$$H_{j,k+\frac{1}{2}}^y = \frac{1}{2} \left( B_{j,k+1} - \Omega_{j,k+\frac{1}{2}}^y \right)^n$$

The nonstandard notation

$$H_{j+\frac{1}{2},k}^x E = \frac{1}{2} \left( A_{j+1,k} E_{j+1,k} - \Omega_{j+\frac{1}{2},k}^x E \right)^n$$

is used and

$$\Omega_{j+\frac{1}{2},k}^x E = R_{j+\frac{1}{2}} \text{diag} \left[ \psi(a_j^{l+1}) \right] R_{j+\frac{1}{2}}^{-1} (E_{j+1,k} - E_{j,k})$$

$$\Omega_{j,k+\frac{1}{2}}^y E = R_{k+\frac{1}{2}} \text{diag} \left[ \psi(a_k^{l+1}) \right] R_{k+\frac{1}{2}}^{-1} (E_{j,k+1} - E_{j,k}).$$

Here $A_{j+1,k}, B_{j,k+1}$ are Jacobians of $F$ and $G$ evaluated at $(j+1,k)$ and $(j,k+1)$. The value $D_{j,k}$ is the Jacobian of the source term $S$ evaluated at $(j,k)$, and $E_{j,k} = U_{j,k}^{n+1} - U_{j,k}^n$. The
expression $\text{diag}(\varepsilon^t)$ denotes a diagonal matrix with diagonal elements $\varepsilon^t$. To compute $(4.4f,g)$, a triple matrix multiplication of dimension $(ns + 3, ns + 3)$ has to be performed at every grid point. For steady-state applications, one can simplify $(4.4f,g)$ as

$$
\Omega^{x}_{j+\frac{1}{2},k} E = M_x I (E_{j+1,k} - E_{j,k})
$$

(4.4h)

$$
\Omega^{y}_{j,k+\frac{1}{2}} E = M_y I (E_{j,k+1} - E_{j,k})
$$

(4.4i)

The scalar values $M_x$ and $M_y$ are

$$
M_x = \max_i \psi(a^l_{j+\frac{1}{2}})
$$

(4.4j)

$$
M_y = \max_i \psi(a^l_{k+\frac{1}{2}})
$$

(4.4k)

and $I$ is the identity matrix. Note that $(4.4h,i)$ involve scalar multiplications only. Other linearized treatments can be found in [31].

4.2 Stiff Source Terms, ADI Approaches and Relaxation Methods

The extra stiff term $D^a_{j,k}$ on the implicit operator $(4.4)$ complicated the solution procedures for the commonly used ADI procedures. Normally, if $D^a_{j,k}$ are not stiff, one can reformulate $(4.4)$ by an ADI procedure like the Beam and Warming [35] algorithm for an efficient solution process. Unfortunately, the $D^a_{j,k}$ considered here are stiff; consequently the additional higher order terms due to the ADI formulation can no longer be ignored. In a different context, Van Dalsem and Steger [36] suggested a remedy if $D^a_{j,k}$ is a diagonal matrix with identical diagonal elements. However, for chemically reacting flows, the matrix $D^a_{j,k}$ is full for the upper $(ns,ns + 3)$ entities and no straightforward way of utilizing ADI approaches for nonlinear system cases with a general stiff source terms can be found.

Recently, Gnoflo, McCandless and Yee [11] have successfully demonstrated the usefulness of a point-relaxation method on the implicit symmetric TVD scheme $(4.4)$ for a loosely coupled chemically nonequilibrium flow. Here a similar point-relaxation or line-relaxation method is proposed for the fully coupled system (3.1). Despite the fact that a larger equation set is involved than in [11], the extra operations are minimized by making use of the simplification procedure of section II. For a point-relaxation method, the size of the matrix inversion for $(4.4)$ is $(ns + 3, ns + 3)$ as opposed to the loosely coupled method of [11], where the size is $(ns,ns)$. The gain in the freedom of controlling the appropriate amount of numerical dissipation to each individual wave more than compensates for the extra expense. More importantly, solving the fully coupled system is believed to have a better convergence rate than the loosely coupled approach. All of the necessary terms required for the implicit scheme $(4.4)$ are derived in section II. The implicit operator of $(4.4a)$ is diagonally dominant for $D_{j,k} = 0$. Therefore one has to make sure that the type of source term to be used will not destroy the diagonally dominant property which is required for some relaxation methods.

For steady-state application, an algorithm utilizing the TVD scheme for the viscous flows is to difference the hyperbolic terms the same way as before, and then central-difference the
Viscous terms. The final algorithm is the same as equation (4.4), except that the spatial central differencing of the viscous term is added to the right-hand side of equation (4.4). Numerical tests, comparison with other approaches, and recommendations will be reported in a future publication.

V. Concluding Remarks

Two numerical algorithms are proposed for hyperbolic conservation laws that are suitable for thermally and chemically nonequilibrium flows in the hypersonic regime. The specific properties of the governing equations for fluid flow of this type were taken into consideration for more efficient solution procedures. The main areas of consideration are to minimize operation counts, decrease truncation errors, increase the allowable time-step constraint imposed by the stiff source terms, and expand the shock-capturing capability beyond classical approaches. Details of all the considerations are described. A preliminary test problem shows certain advantages of the proposed semi-implicit high-resolution shock-capturing scheme over the classical ways of supplying numerical dissipation. More numerical testing and study will be pursued in the immediate future.

Acknowledgement

The authors wish to thank K. Uenishi and A. Kumar for the kind permission to use their computer code as a base code for the development of the current research code in verifying the applicability of the proposed numerical method to chemically reacting flows. The first author also wishes to thank M. Vinokur, E. Murman, G. Carofano, and R. F. Warming for many stimulating discussions. In particular, thanks to M. Vinokur for guidance in understanding the physics of gas dynamics.

References


Fig. 1. Schematic of physical problem.
Fig. 2. Density contours of a moving planar shock wave over an obstacle.
$p = 1 \text{ atm}$
$T = 1200 \text{ K}$
$M = 4$

Fig. 3. Geometry and inflow condition.
Fig. 5. Species distribution along \( z = 0.13 \) cm.
Some numerical aspects of finite-difference algorithms for nonlinear multidimensional hyperbolic conservation laws with stiff nonhomogeneous (source) terms are discussed. If the stiffness is entirely dominated by the source term, a semi-implicit shock-capturing method is proposed provided that the Jacobian of the source terms possesses certain properties. The proposed semi-implicit method can be viewed as a variant of the Bussing and Murman point-implicit scheme with a more appropriate numerical dissipation for the computation of strong shock waves. However, if the stiffness is not solely dominated by the source terms, a fully implicit method would be a better choice. The situation is complicated by problems that are higher than one dimension, and the presence of stiff source terms further complicates the solution procedures for alternating direction implicit (ADI) methods. Several alternatives are discussed. The primary motivation for constructing these schemes was to address thermally and chemically nonequilibrium flows in the hypersonic regime. Due to the unique structure of the eigenvalues and eigenvectors for fluid flows of this type, the computation can be simplified, thus providing a more efficient solution procedure than one might have anticipated. In a companion paper, extension and application of the semi-implicit method to three-dimensional chemically nonequilibrium flows in generalized coordinates will be presented. A preliminary numerical computation for a five-species reacting flow is shown to be oscillation-free around the shock.